14 Evolutionary Constraints on Emergence of Plant RNA Viruses

Santiago F. Elena

Abstract

Over the recent years, agricultural activity in many regions has been compromised by a succession of devastating epidemics caused by new viruses that switched host species, or by new variants of classic viruses that acquired new virulence factors or changed their epidemiological patterns. Although viral emergence has been classically associated with ecological change or with agronomical practices that brought in contact reservoirs and crop species, it has become obvious that the picture is much more complex, and results from an evolutionary process in which the main players are the changes in ecological factors, the tremendous genetic plasticity of viruses, the several host factors required for virus replication, and a strong stochastic component. The present chapter puts emergence of RNA viruses into the framework of evolutionary genetics and reviews the basic notions necessary to understand emergence, stressing that viral emergence begins with a stochastic process that involves the transmission of a pre-existing viral strain with the right genetic background into a new host species, followed by adaptation to the new host during the early stages of infection.

Introduction: what is an emerging virus?

Which viruses deserve the qualification of 'emerging' is somewhat controversial. The word is frequently used to describe the appearance of a hitherto unrecognized viral infection or a previously recognized one that has expanded into a new ecological niche or geographical zone, often accompanied by a significant increase in symptom severity (Cleaveland *et al.*, 2007). According to the USA Center for Disease Control and Prevention, an emergent virus should meet the following definition: a disease of infectious origin whose incidence has increased within the past decades or threatens to increase in the near future. However, this definition is somewhat vague and misleading, and a virus may be classified as emerging for reasons that have little to do with the spirit of the term *emerging*, such as increasing awareness, the adoption of improved diagnostic tools, or the discovery of previously uncharacterized agents for already known diseases. Similarly, truly emerging viruses may not be recognized as such due to poor case reporting, or difficulties in diagnosis. Following Woolhouse and Dye (2001), a more rigorous definition of an emerging virus would be the causal agent of 'an infectious disease whose incidence is increasing following its first introduction into a new host population or whose incidence is increasing in an existing host population as a result of long-term changes in its underlying epidemiology'. This definition implies that the virus is spreading in the host population upon its first description and it has nothing to do with changes in symptomatology. According to Woolhouse and Dye's definition, the epidemic spread during the late 1980s and early 1990s of necrogenic strains of cucumber mosaic virus (CMV) on tomato crops in eastern Spain (Escriu *et al.*, 2000) would hardly be considered as an emerging virus. However, it would be qualified as an emerging disease by Cleaveland's definition. By contrast, pepino mosaic virus (PepMV), which

was first described infecting tomatoes in 1999 in The Netherlands (Van der Vlugt *et al.*, 2000), and is now quickly spreading across Europe and beyond, should be considered as a paradigm of emerging viral infection by Woolhouse and Dye's definition.

However, I find that no definition is entirely satisfactory, and the discrepancy entirely semantic, and thus hereafter I will use a slight modification of Woolhouse and Dye's definition that incorporates also changes in pathology. This will allow me to classify both of the above examples as emerging plant diseases.

The sources of emerging viruses are different host species, the reservoirs, in which the virus is already established. Species jumps (aka spillovers) have given rise to devastating epidemics in crop species. However, there are numerous examples of species jumps that have had far less dramatic consequences (examples are cotton leaf curl virus infecting ancient cotton cultivars in India, and maize rough dwarf virus infecting maize in the Mediterranean region before the introduction of the American high-yield hybrid cultivars – see Thresh (2006) for a review) and there are even many viruses that have a long history of routinely jumping between species without triggering major epidemics (e.g. CMV).

In the following sections I will go through the mechanisms and processes that are behind plant RNA virus emergence. These processes will be divided into three phases. The first phase accounts for the mechanisms and limitations for jumping the species barrier. The second phase includes the study of the evolutionary dynamics that end up with a virus well adapted to its new host. The third phase comprises the epidemiological spread of this well-adapted virus in the new host population.

I will focus this review entirely on RNA viruses because of their apparent larger evolvability, the consequence of combining highly error-prone replication, large population sizes and rapid replication rates (Elena and Sanjuán 2008). For the moment, let's reserve the discussion on whether RNA viruses are more evolvable than DNA ones for a different place, and let's assume that the principles that drive RNA virus emergence will not be substantially different from those driving DNA virus emergence (Chapter 15). By doing so, whatever lesson may be taken from this review may help readers to understand the emergence of their favourite plant DNA virus.

Phase I of viral emergence: ecological determinants of cross-species spillovers

The first step in virus emergence is the exposure of the new host species to the virus (Fig. 14.1). The rate of exposure will be a function of the ecology and behaviour of the two hosts, and of the transmission biology of the virus, including any relevant vector that may be involved.

Ecological disturbance and geographical distribution of alternative hosts

Contact between donor and recipient hosts is a precondition for virus spillovers, and it therefore depends on the ecology, biogeography and behavioural separation between reservoir and recipient species. Factors that affect the geographical distribution of hosts (e.g. trading of ornamental plants, the introduction of uncontrolled cultivars, or the conversion of wild tropical areas to cultivable) or that facilitate the spread of vectors, tend to promote viral emergence. Jones (2009) has identified up to nine different scenarios for emergence upon which introduced plants are exposed for the first time to indigenous viruses and vectors associated with the native flora. These scenarios represent situations in which the donor and recipient hosts, the vector and the virus may interplay, and involve jumps from the native flora to the introduced crop and *vice versa*.

The density of the recipient host population is important in the onward transmission and epidemic potential of any transferred virus (Woolhouse *et al.*, 2005). Therefore, agricultural intensification and extensification strongly facilitates the establishment and epidemic spread of emerging viruses.

The ongoing global warming will also unavoidably affect the rate at which emergent plant viruses arise. For instance, small changes in average temperature can suffice to produce significant shifts

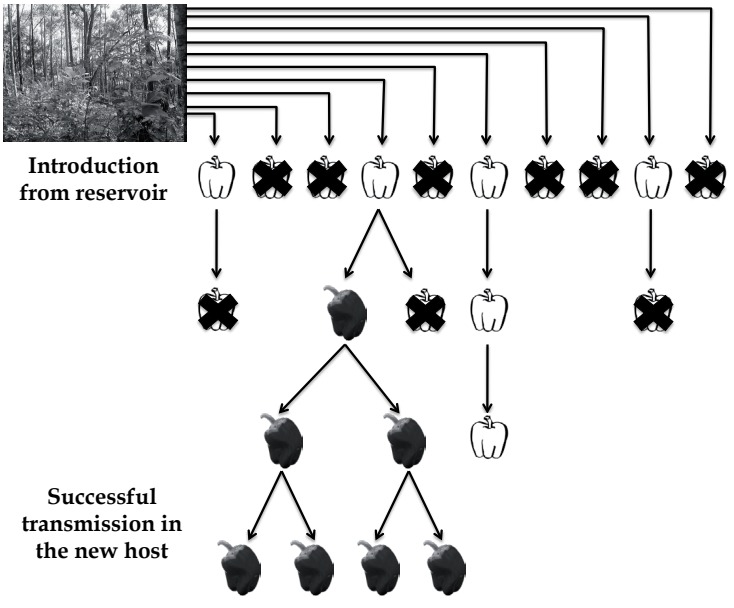

Figure 14.1 Host-switching process. In a first phase, the virus jumps from its natural host to the new one. Each arrow departing from the forest picture and ending in the pepper crop represents an independent spillover. Most of these transmissions will not produce a successful infection (black crosses). In a few cases, the virus will replicate enough to be transmitted for a second time or even a third time, but without triggering an epidemic. In a very few instances, the virus will have increased its fitness in the new host, allowing successful transmissions, leading to it becoming epidemic (HERE INDICATED BY THE CHANGE IN COLOUR **[IS THIS OK AS FIGURE IS IN BLACK AND WHITE?]**).

in the distribution and abundance of arthropod vectors, with increases in their numbers leading to higher risk of epidemics (Anderson *et al.*, 2004). Similarly, soil-borne viruses transmitted by fungi or nematodes may have a higher incidence in temperate zones as they become warmer, but by contrast, they will gradually be less important in tropical zones as they become drier and make the spread of the zoospores more difficult. Temperature increases will also directly affect the frequency and duration of virus epidemics, since virus replication rate is affected by temperature, but also because plant defences are also temperature-dependent (e.g. the RNA silencing pathway), and they may become ineffective. Similarly, the effectiveness of pathogen-induced resistance genes is influenced by CO_2 and O_3 (Garrett *et al.*, 2006). All these environmental alterations will change the selective pressures operating on plant viruses, thereby affecting their rates of evolution and the likelihood of emergence.

Levels of genetic variability of viruses in their reservoir hosts: mutation and recombination rates

The initial infection of individuals of the new host species is a pivotal step in viral emergence. However, most viruses transferred to new hosts are not adapted; they replicate poorly and are inefficiently transmitted. Therefore, the preexistence of host-range mutants among the standing natural variation in the reservoir host increases the probabilities of a successful jump. The amount of standing variation would depend, in a first instance, on the rates of mutation and recombination, in a second instance on the distribution of mutational effects on fitness (that is, the selective value of each mutation generated), and in a third instance on the strength of genetic drift and gene flow among demes. Therefore, a first fundamental question that I will try to answer in this section is what are the values of these two rates in plant viruses.

Table 14.1 shows estimates of mutation rates obtained for CMV, cowpea chlorotic mottle virus (CCMV), chrysanthemum chlorotic mottle viroid (CChMVd), tobacco etch virus (TEV), tobacco mosaic virus (TMV), and wheat streak mosaic virus (WSMV) on different hosts. It is important to note that all values shown in Table 14.1 were estimated by evaluating the genetic variability present in plants infected with inocula containing no genetic variability. Therefore, estimates correspond to the upper bound of possible values (Sanjuán *et al.*, 2009), except in the case of CChMVd, where only lethal mutations were taken into consideration for the computation (Gago *et al.*, 2009). The first conclusion that can be drawn from Table 14.1 is that heterogeneity exists among different viruses in their mutation rates, with values ranging over almost two orders of magnitude ($0.2–17 \times 10^{-4}$ substitutions per site and generation). This broad range of mutation rates is in the same ballpark as estimates obtained for animal viruses. A second interesting observation from Table 14.1 is that, for a given virus, the mutation rate strongly depends on the host in which it was estimated, with differences being as large as 70-fold for TMV.

A fact that is usually not taken into consideration is that, for a given mutation rate, the actual number of mutations per genome per cell strongly depends on whether replication occurs according to Luria's stamping machine or geometrically. If replication follows a stamping machine model, the number of mutations will be smaller than if replication occurs geometrically (Sardanyés *et al.*, 2009). This is intuitively easy to understand: a stamping machine always replicates the same template, and therefore mutations appear in a mutation-free background, whereas geometric replication implies that offspring molecules can serve as templates for further rounds of replication and, thus, mutations may appear in an already mutated genome. Despite its importance, not much evidence exists on what is the exact mechanism of replication for plant viruses.

Because recombination is a process that potentially increases fitness by creating advantageous genotypes and removing deleterious mutations, it might be supposed that it bolsters the process of emergence. However, this possibility is still controversial. While some authors have proclaimed that it may assist the process of cross-species transmission (Chare and Holmes 2006; Codoñer and Elena 2008), others have pointed out that the association between recombination and emergence is circumstantial (Holmes 2008). To get an idea of the impact of recombination in plant RNA viruses, I searched for 'plant RNA virus recombination' in PubMed. Over 560 references were retrieved that illustrate examples of recombinant genotypes among plant viruses. However, only one of these studies is reporting an estimate of the recombination rate *in vivo*. In all other cases, reports are based in the analyses of epidemiological sequence data. These phylogenetic data, although very illustrative, have at least one major drawback: they only inform about successful recombinant genotypes sorted out by natural selection and that generally induce new pathologies; thus they may underestimate the real recombination rate.

Chare and Holmes (2006) made an extensive phylogenetic analysis of recombination in plant viruses. They analysed 36 virus species belonging to six families and found compelling evidences of recombination in one third of these viruses, also confirming that the frequency of recombinants differed widely among families, with the potyviruses showing higher frequencies than the other families. A higher frequency of recombinant genotypes does not means that potyviruses are more recombinogenic than the other species. At face value, the observation only means that recombinant genotypes have increased their frequency in populations due to some selective advantage.

The only report available for an *in vivo* recombination rate was obtained for the dsDNA pararetrovirus cauliflower mosaic virus (CaMV) by Froissart *et al.* (2005). These authors found that half of the CaMV genotypes sequenced were recombinant, assuming that replication occurs geometrically (which may be not entirely the case), the authors calculated a recombination rate in the range $2–4 \times 10^{-5}$ per base and replication cycle, of the same order of magnitude as the estimates for mutation rates shown in Table 14.1.

Table 14.1 Upper-limit estimate for the mutation rate for several plant RNA viruses and a viroid on different hosts

Virus	Host	Rate (± SEM) × 10^{-4}	Reference
CMV	*Capsicum annuum*	15.34 ±0.71	Schneider *et al.* (2001)
	C. annuum	1.39 ± 0.07	Pita *et al.* (2007)
	Nicotiana benthamiana	6.64 ± 0.95	Schneider *et al.* (2000)
	Nicotiana tabacum	0.20 ± 0.09	Pita *et al.* (2007)
CCMV	*N. benthamiana*	5.29 ± 4.93	Schneider *et al.* (2000)
CChMVd	*Dendranthema grandiflora*	25.00 ± 6.00	Gago *et al.* (2009)
TEV	*N. tabacum*	0.30 ± 0.03	Sanjuán *et al.* (2009)
TMV	*C. annuum*	11.02 ± 0.12	Schneider *et al.* (2001)
	Collinsia heterophylla	4.74	Kearney *et al.* (1999)
	Fagopyrus esculentum	4.55	Kearney *et al.* (1999)
	Lycopersicum esculentum	1.45 ± 0.51	Schneider *et al.* (2001)
	N. benthamiana	4.21 ± 0.69	Schneider *et al.* (2000)
	N. tabacum	4.14	Kearney *et al.* (1999)
	N. tabacum	0.24 ± 0.00	Malpica *et al.* (2002)
	Phacelia campanularia	16.81	Kearney *et al.* (1999)
	Plantago sp.	8.50	Kearney *et al.* (1999)
	Solanum nigrum	4.21	Kearney *et al.* (1999)
	Tagetes erecta	8.15	Kearney *et al.* (1999)
WSMV	*Zea mays*	9.01 ± 0.90	Hall *et al.* (2001b)

This coincidence suggests that recombination should be a source of variation as important as mutation for viral emergence. However, given the differences in genomic architecture of CaMV and RNA viruses, caution needs to be expressed against generalizing this number to plant RNA viruses until more empirical estimations become available.

Recombination rates in plant RNA viruses are controlled by two factors: the ability of the virus in question to undergo template switching and the multiplicity of infection (MOI) during disease progression. The first factor would clearly vary among viruses as a function of their biology, and for example, negative-strand RNA viruses are expected to be less recombinogenic, since their RNA is never naked (Chare *et al.*, 2003). The second factor, namely the frequency at which a cell is infected with at least two different viral genomes, would likely depend on the peculiarities of each virus–host pair, and has received little attention yet. In a groundbreaking report, González-Jara *et al.* (2009) undertook the task of evaluating the frequency of multiple infections within a single infected host for TMV in *Nicotiana benthamiana*. These authors tracked the kinetics of infection of two different TMV genotypes (respectively labelled with RFP and GFP) by counting the number of cells singly and co-infected. Their results suggest that MOI is high during infection, although the value decreased as the infection progressed, both in inoculated and systemically infected leaves. This decline in MOI opens the possibility for the existence of superinfection-inhibition mechanisms in TMV, but this point clearly needs empirical support. The results just described for TMV may not be general, since they contrast somewhat with the results of a study undertaken with several labelled potyviruses; during mixed infections with different genotypes of the same potyvirus, they exclude each other, whereas two different potyviruses can

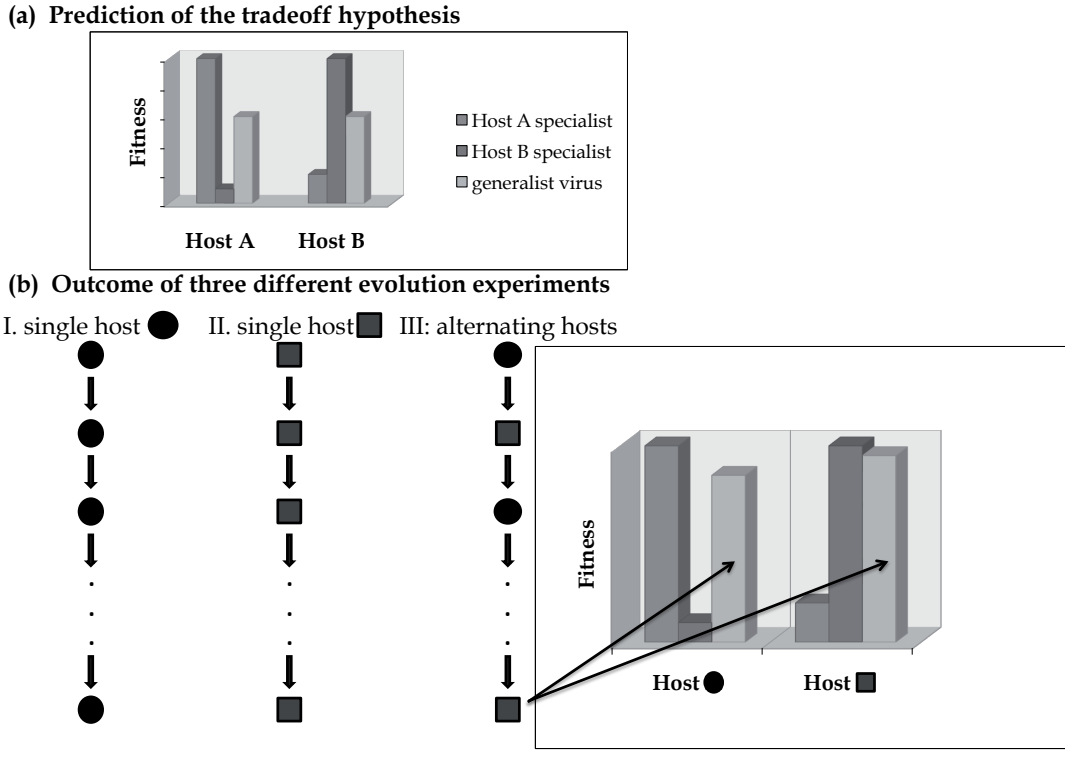

Figure 14.2 Fitness trade-offs across hosts. (a) Expected fitness for specialist and generalist viruses if a trade-off exists. Although both specialist genotypes perform well in their respective hosts, each is poorly adapted in the other host. The light grey bars illustrate the behaviour of a generalist virus that performs fairly well in both hosts, but has lower fitness than either specialist in its preferred host. According with this picture, a specialist virus will always outcompete a generalist on its host, but if hosts vary in time or space, the generalist may have an overall advantage. (b) Outcome of three evolution experiments. Viruses evolved in a single host become specialists on their respective hosts; by contrast, viruses evolved in a fluctuating host landscape become generalists, and improve fitness in both hosts at the same time (light grey bars).

coinfect the same cell (Dietrich and Maiss 2003). Hence, MOI may not be as high in potyviruses as it appears to be in TMV, but the potyvirus results support the existence of some superinfection-inhibition mechanism.

Fitness trade-offs across hosts

A fundamental challenge for host-switching viruses that require adaptation to their new hosts is that mutations that optimize the ability of a virus to infect a new host will likely reduce its fitness in the reservoir (Fig. 14.2). The nature of these fitness trade-offs and how they affect cross-species transmission is an important and active area of research. We have recently written a review article on this topic, illustrated with many examples from plant, animal and bacterial viruses (Elena *et al.*, 2009), hence I will not repeat the same information here. Readers interested in the details can check the review article and the references therein. Here I will just provide a short overview of the topic.

By specializing in a single host, viruses may reduce interspecific competition at the cost of accessing a more limited set of available resources. In stark contrast, the advantages of generalism are more obvious: a generalist virus would be able to exploit multiple hosts, thus enhancing its fitness. Since generalist plant viruses are not the norm (Malpica *et al.*, 2006), it is generally assumed that

generalism comes with a cost, in keeping with the adage that a 'jack-of-all-trades' is a master of none. It has been suggested that evolution should favour specialists, because there are trade-offs that limit the fitness of generalists in any of the alternative hosts, or because evolution proceeds faster with narrower niches (Fig. 14.2a). Fitness trade-offs can be generated by different mechanisms, antagonistic pleiotropy being the simplest and most intuitive one. Antagonistic pleiotropy means that mutations that are beneficial in one host may be deleterious in an alternative one. A second mechanism that promotes trade-offs is mutation accumulation, in which neutral mutations accumulate by drift in genes that are useless in the current host but may be essential in a future new one. Although both mechanisms involve differences in mutational fitness effects across hosts, it is necessary to stress that they are by no means equivalent phenomena; while natural selection is the only reason for the trade-off in the former mechanism, genetic drift is important in the latter.

Much experimental evidence suggests that whenever a virus switches hosts, acquiring the ability to replicate in a new host imposes a fitness burden in the original host. This may be a consequence of the different selective requirements characteristic of different hosts (Fig. 14.2b). However, some evidence also suggest that the fitness of a virus simultaneously facing multiple hosts is either constrained by the most restrictive one, or there is no trade-off at all (Fig. 14.2b). In this respect, the extent to which generalism evolves depends on the frequency at which viruses transmit among heterologous hosts (Wilke *et al.*, 2006). When transmission among heterologous hosts represents an infrequent event, the viral population essentially adapts to the current host. However, if heterologous transmissions are frequent, the viral population behaves as if the fitness landscape did not change at all, but was the average of the changing landscapes (Wilke *et al.*, 2006). The behaviour at intermediate oscillation frequencies rests between these two extremes.

What are the causes for fitness trade-off across hosts? Most of the accumulated evidence suggests that antagonistic pleiotropy is the principal, although certainly not the only reason (Elena *et al.*, 2009). Antagonistic pleiotropy may be an unavoidable consequence of the small size of viral genomes, which in many instances contain overlapping genes and encode multifunctional proteins, making it extremely difficult to optimize one function without jeopardizing another.

Genetic relatedness between reservoir and naive hosts

The next question that pops up is whether some viruses are more able to jump species barriers than others. A compelling idea in this respect is that there are phylogenetic constraints to this process, such that the more closely related the reservoir and the new host, the greater the chances for a successful spillover (DeFilippis and Villareal 2000). There are good mechanistic reasons to believe that a relationship exists between host's phylogenetic distance and the likelihood of viral emergence. It can be argued that if the ability to recognize and infect a host cell is important for cross-species transmission, then phylogenetically related species are more likely to share related cell receptors and defence pathways. However, others support the view that spillovers have occurred between hosts that can be either closely or distantly related, and no rule appears to predict the susceptibility of a new host (Holmes and Drummond, 2007).

Whether or not genetic relatedness between reservoir and new hosts may be a factor for host switching, the rate and intensity of contact may be even more critical. Viral host switches between closely related species (e.g. species within the same genera) may also be limited by cross-immunity to related pathogens (Parrish *et al.*, 2008). Or using the words of Holmes and Drummond (2007) 'although a species might be exposed to a novel pathogen, they might, through a combination of shared common ancestry and good fortune, already posses a sufficient immune response to prevent the infection from being established'.

Phase II of viral emergence: adaptation to the new host

In the previous section, I reviewed what factors may make certain RNA viruses more prone to emergence than others. In addition to ecological factors and the genetic relatedness between

reservoir and naïve host species, I have put the emphasis on the virus genetic factors that determine the presence of abundant genetic variability in the source viral population and in the likelihood that this population may contain genetic variants with the ability of infecting and replicating, to some extent, in the putative new host. In the following sections I will move one step ahead and discuss some of the factors that may determine the adaptation of the emerging virus to its new host.

Factors modulating within-host adaptation dynamics: effect and distribution of mutational effects and of epistasis

The evolutionary fate of a population in a constant environment depends on the distribution of mutational effects on fitness. This is the fraction of all possible mutations that are beneficial, neutral, deleterious, or lethal. For a well-adapted virus and given the compactness of viral genomes, with many cases of overlapping genes and multifunctional proteins, most mutations are expected to fall into the deleterious and lethal categories. However, the distribution of fitness effects on a given genotype are rarely constant across environments, and the contribution of each category to the overall fitness will vary widely, depending on the overlap between the alternative environmental conditions (Martin and Lenormand 2006). This environment-dependence of the distribution of mutational effects may impact the likelihood of adaptation of a virus after host switching. For instance, if the environment provides new opportunities for the virus, the fraction of beneficial mutations may be increased either by moving the average of the distribution towards more positive values while keeping the shape constant (Fig. 14.3), or alternatively, without affecting the mean but increasing the variance (Fig. 14.3). So far, the information in hand for making educated guesses about the environmental effect is scarce beyond a few model organisms (e.g. *Escherichia coli* and *Saccharomyces cerevisiae*), and certainly inexistent for plant RNA viruses. In a recent study (Carrasco *et al.*, 2007), we explored the distribution of single-nucleotide substitution mutational effects for TEV on its natural host, tobacco. In short, we found that most mutations were strongly deleterious for the virus, with up to 41% of mutations being lethal, 36% significantly deleterious (on average reducing fitness 41%), 23% had no measurable effect on fitness (i.e. they were neutral on tobacco), and no beneficial mutations were detected, not surprisingly, in the natural host. It is relevant that these results are in good qualitative agreement to other reports for animal viruses (Sanjuán *et al.*, 2004a) and bacteriophages (Domingo-Calap *et al.*, 2009), and taken all together draw a picture showing viral RNA genomes as very sensitive to mutational effects. Characterizing the distribution of mutational effects across a panel of possible alternative hosts, varying in genetic relatedness to the natural one, is a very important task.

However, given the high mutation rate of RNA viruses, mutations may not appear as single events but genomes may contain multiple hits. Determining the way in which mutations interact in determining viral fitness is another important issue that, for example, determines whether certain evolutionary pathways (i.e. genetic combinations) are more likely than others, which indeed determines the ruggedness of the fitness landscapes wherein viral populations move (Weinreich *et al.*, 2005). If mutational effects are always additive, the shape of the landscape will be smooth, with a single peak emerging from a flat surface (Fujiyama-like landscape). By contrast, the more the average interaction deviates from additive effects, the more fitness peaks of different heights may exist in a landscape (Alps-like landscape). Unfortunately, a direct evaluation of the extent and intensity of epistasis in the genome of plant RNA viruses is not yet available, and we can only guess that the dominant type of epistatic interaction in these viruses would be similar to that observed for animal viruses (Bonhoeffer *et al.*, 2004; Sanjuán *et al.*, 2004b) and bacteriophages (Burch *et al.*, 2004; Rokyta *et al.*, 2005). Information from these other systems suggests that, on average, mutations in viral genomes interact in a negative way, that is, the observed effect of two mutations together is lower than expected from their individual effects. This diminishing-returns effect is expected to speed up the rate of adaptation (Sanjuán *et al.*, 2005). Similarly to

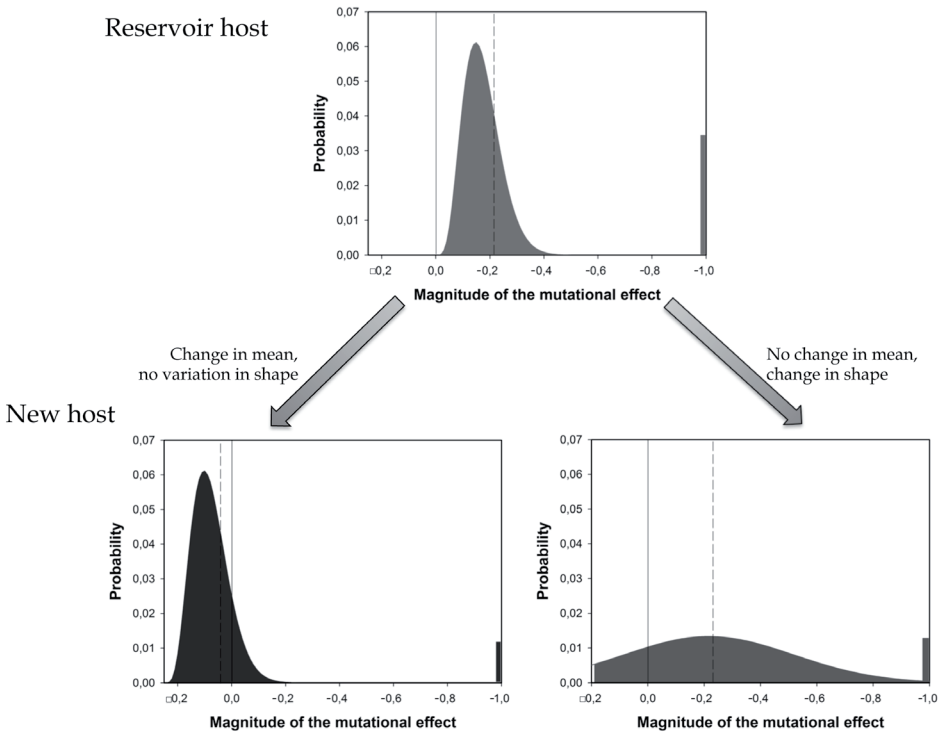

Figure 14.3 Possible effects of host switching on the distribution of mutational effects on viral fitness. In all cases, the average mutational effect is indicated by the dashed vertical line, and the neutral case by the solid vertical line. The area under the curve to the left of the solid vertical line represents the fraction of beneficial mutations, whereas the area to the right of this line represents the fraction of mutations with deleterious effects. The upper diagram shows the distribution of mutational effects on the reservoir host. The lower diagrams show two potential host effects: the left one represents a change in the mean effect without affecting the shape of the distribution, the right figure represents a change in shape without altering the average value. In both cases the fraction of mutations with beneficial effects is increased.

what was mentioned above for single mutations, the cause of negative epistasis may be found in the existence of overlapping genes in RNA genomes encoding for multifunctional proteins (Elena *et al.*, 2006).

If nothing is known about host's effects on the distribution of mutational effects on viral fitness, even less is known about its effect on epistasis. This information is critical for understanding plant RNA virus emergence.

Evasion from host defences

After host switching, it is critical for the virus to deal with the plant defence mechanisms. Plants have a wide variety of complex responses to viral infection, including non-specific resistance mechanisms, both innate and acquired (e.g. hypersensitive and ROS responses) and specific (e.g. gene-for-gene, systemic acquired resistance –SAR- and RNA silencing). All these forms of resistance have been reviewed recently by Jones and Dangl (2006), by Király *et al.* (2007) and are described in several other chapters (Chapters 8, 9, 10 and 11). Therefore, I will not extend myself here discussing all possible evolutionary solutions that viruses may find to escape from each mechanism. In contrast, I will just comment on one that I find particularly interesting from an evolutionary perspective because of its conservation across kingdoms: RNA silencing.

Because its properties of memory and sequence specificity are similar to those of vertebrate's immune system, one of the mechanisms that has attracted more attention during the last

decade is virus-induced RNA silencing (Voinnet 2001; Waterhouse *et al.*, 2001; see also Chapter 6). Not surprisingly, soon after the identification of RNA silencing as a plant response to viral infection, the existence of viral proteins with the capacity of interacting with different components of the silencing pathway, blocking the antiviral response and enhancing virus accumulation and systemic movement was reported (reviewed in Li and Ding 2006; Díaz-Pendón and Ding 2008; see also Chapter 7). The evolutionary implications of these suppressor proteins has not been fully explored yet, but in a recent compensatory evolution experiment we have shown that the TEV suppressor protein HC-Pro may be under strong stabilizing selection, suggesting that it is detrimental for the virus both to reduce and to increase the strength of suppression (Torres-Barceló *et al.*, 2009).

In addition to the evolution of active siRNA evasion mechanisms, the high mutation rate of plant RNA viruses may also facilitate evasion from RNA silencing by generating escape mutants at a high rate. To evaluate the likelihood of generating mutants able of escaping from the selective pressure imposed by a single siRNA, Lin *et al.* (2009) inserted a non-coding sequence into the genome of turnip mosaic virus (TuMV). This non-coding sequence was targeted by an artificial microRNA transgenically expressed by the host plant *N. benthamiana*. As expected, transgenic plants were resistant to TuMV infection. Then, each of the 21 nt in the siRNA target sequence was mutated and the pathogenicity of each single-nucleotide substitution mutant evaluated in the transgenic plants. Mutations at six positions in the target rendered viruses with high pathogenicity, most of these mutations being located at the 5′ end of the siRNA; mutations at nine positions scattered along the siRNA sequence only produced a minor increase in pathogenicity. Nonetheless, the presence of mutations at any site in the target sequence allowed the mutant virus to replicate enough to produce additional mutations that further increased the pathogenicity of the mutant virus (Lin *et al.*, 2009). This experiment serves as example of the easiness by which a population of RNA viruses may escape from the surveillance of siRNAs simply by mutation. However, it is worth noting that (i) in a more realistic situation multiple siRNAs are produced against the viral genome, and (ii) the target sequence encodes a protein, implying that not all changes would be equally permitted due to their fitness consequences.

I do not want to close this section without mentioning that in a recent study we found evidence suggesting that during the adaptation of TEV to the non-natural host *Arabidopsis thaliana*, the expression pattern of genes involved in stress responses (including SAR and RNA silencing) were significantly downregulated to the same level as was measured in the mock-inoculated plants (Agudelo-Romero *et al.*, 2008b). These stress genes were all significantly upregulated in plants infected with the ancestral non-adapted virus (Agudelo-Romero *et al.*, 2008a). If confirmed, this result would suggest that one way that natural selection might find to optimize viral fitness in a novel host is by making it undetectable by plant defences. A final implication of these results is to call for extra precaution when reading the results reported by several authors on changes in gene expression in control versus virus-infected *Arabidopsis* plants. Almost in every case, the viruses employed for infecting *Arabidopsis* were not previously adapted to this artificial host. If adaptation changes the way the virus interacts with the plant, then these experiments may inform us of nothing beyond what may be a general response to stress.

Metapopulation dynamics within infected hosts

Plant architecture creates a spatially structured environment for plant viruses. This means that the viral population replicating within an infected plant cannot be considered as a single panmictic population, but as a collection of subpopulations each replicating in different leaves. Spatial structure imposes strong conditions on the spread of beneficial mutations that may improve the fitness of an emerging virus on its new host. Spatial structure exists at different levels: from leaves to branches.

Using plum pox virus (PPV) clones labelled with two different flavours of fluorescent protein, Dietrich and Maiss (2003) were able to observe

that the two populations excluded each other during the colonization of *N. benthamiana* epidermal cells. Only a minority of cells in the contact region between growing foci were doubly infected. This spatial separation reduces the opportunities for competence between genetic variants, thus reducing the efficiency with which natural selection may increase overall population fitness. Furthermore, this strong spatial structure imposes a barrier on the fixation of beneficial mutations in the whole metapopulation, regardless of the magnitude of their beneficial effect, if they appear in cells that are already confined by cells infected with other viral genotypes.

Certainly not the only one, but for me the clearest demonstrations that viral populations differentiate into genetically isolated subpopulations within a single plant was reported by Jridi *et al.* (2006) for PPV. These authors analysed the population structure of PPV within a single infected *Prunus persica* tree 13 years after inoculation. They observed that following the systemic invasion of the host, the virus population differentiated into several subpopulations that were isolated in different branches. These subpopulations subsequently differentiated into other subpopulations, with little to no genetic exchange between distal parts. Very nicely, the phylogenetic tree linking PPV genomes isolated from different leaves and branches matched the branching pattern of the tree.

One may ask whether this segregation of viral populations into different subpopulations is driven by fitness differences, or if the determination of the genotype colonizing a distal tissue is a purely stochastic process. In recent years, different groups had been engaged in estimating the strength of population bottlenecks during the colonization of distal tissues. The standard population genetic parameter used to this end is the effective population size (N_e). Hall *et al.* (2001a) used a simple experimental design to estimate N_e during systemic colonization of WSMV. In short, they mixed two different strains of WSMV and used the mixture to coinfect wheat seedling. Then, they determined how many tillers were infected with a single strain versus how many were coinfected. The frequency data were then fitted to a Binomial distribution and determined that N_e for systemic colonization was 3–5 genomes. Sacristán *et al.* (2003) used a similar co-inoculation approach and estimated that during systemic colonization of new leaves by TMV, the size of the founder N_e was in the order of units. In a rather similar experiment that involved 12 genetic markers, Li and Roossinck (2004) showed that the genetic variance of CMV populations replicating in a single leaf was significantly and reproducibly reduced in systemic leaves, with the number of markers present in the systemic leaves ranging between 4 and 8. Unfortunately the authors did not perform any statistical analyses of the data, in order to provide a quantitative value for the expected N_e. Nonetheless, I took my time and used the variance components method described in Monsion *et al.* (2008) to estimate that N_e in these experiments ranged between 12 and 220 genomes. Finally, Monsion *et al.* (2008) estimated, again using a similar experimental design involving six markers, that N_e for CaMV infecting systemic leafs of *Brassica rapa* was in the range of several hundred genomes. In conclusion, N_e estimates widely differ among different viruses. Whether these differences are relevant and the consequence of biological differences among the four viruses studied or an experimental and/or analytical artefact needs to be considered further.

A last consideration I would like to make about the spatial spread of genetic variants is that at high MOI, complementation between genetic variants may slow down the rate at which a beneficial mutation spreads in the population (Frank 2001). When many viral genotypes infect the same host cell, the effective ploidy of the genetic system is high, diluting the contribution of each locus to the phenotype and weakening the selective intensity on each locus. Weaker selection allows maintenance of greater genetic diversity in the population, allowing otherwise deleterious alleles to persist for long periods of time. In such a situation, a genetic system that may avoid superinfection would become beneficial at the long run, by speeding up the rate of evolution at linked loci. This possibility gives further likelihood to the suggestion of González-Jara *et al.* (2009) mentioned above about such mechanisms operating in TMV.

The effect of coinfection with other viruses

I just mentioned that coinfection with genetic variants carrying beneficial mutations and others carrying deleterious alleles may slow down the rate of adaptive evolution. However, at the first stages of Phase II, coinfection between an emerging virus and a different, well adapted virus, may turn out to be beneficial to the former. There are two relevant questions here: first, how often are plants coinfected by more than one virus, and second, do coinfecting viruses share resources?

Interspecific coinfection is commonplace, and the literature is full of references describing the result of coinfection between viruses. In an exhaustive analysis of the incidence of five virus species across 21 species of wild plants, Malpica *et al.* (2006) found that the prevalence of different viruses was not independent of each other, but certain viruses were found together more often than would expected just by chance. In the individual host, coinfection may have variable consequences, ranging from symptom amelioration to synergistic exacerbation (Hammond *et al.*, 1999). Mixed infections can also modify viral traits such as host range (Guerini and Murphy 1999; Hacker and Fowler 2000; García-Cano *et al.*, 2006), transmission rate (Wintermantel *et al.*, 2008), cellular tropism (Moreno *et al.*, 1997; Sánchez-Navarro *et al.*, 2006), or the amount of virus accumulation (Martín and Elena, 2009). Most studies have focused on synergic diseases caused by two ssDNA or ssRNA viruses, particularly by a potyvirus and another ssRNA virus. In many instances, the titre of the non-potyvirus increases, while that of the potyvirus is not altered; this enhancement being explained by the RNA silencing suppression activity of the potyviral HC-Pro (Dunoyer and Voinnet 2005). Nevertheless, these interactions do not always produce synergic diseases, and depending on the particular combination of virus species, accumulation of the counterpart can also decrease (Kokkinos and Clark 2006).

In the previous section, I reviewed evidence that two isolates from the same virus may exclude each other from the same cell, thus creating non-overlapping spatial patterns of genotypes. At least for potyviruses, the exclusion found by Dietrich and Maiss (2003) was limited to PPV variants, while potyviruses belonging to different species did not excluded each other and were found coinfecting the same cells. Given that sequence similarity may still be significant between two members of the same family, coinfection opens the possibility for interspecific recombination or reassortment, and thus the generation of new viral species.

Phase III of viral emergence: epidemiological dynamics

So far, I have been focusing this chapter on the processes that generate genetic variability, as a pre-requisite for emergence (phase I), and the factors that may condition the adaptation of an emerging virus to its new host (phase II). Still, I need to mention, although certainly I will do it very briefly, what characterizes phase III of viral emergence, that is, the epidemiological spread of the new virus in the new host population. Surely, some readers may find the distinction between phases II and III somewhat artificial. I must agree: adaptation to the new host may go hand in hand with the spread in the new host population; the more infections occur, the more likely that beneficial mutations may appear in the viral population, and thus the more likely the viral fitness will be fine-tuned by natural selection.

The basic reproductive ratio and the conditions for an epidemic spread

The epidemiological theory of infectious diseases has a strong theoretical basis, particularly developed to study the spread of infection through a host population (Woolhouse *et al.*, 2005). How big or small an outbreak may be depends on two factors: (i) the number of introduction events, that is, how often the virus spills over from the host reservoir to the new host and (ii) the potential for transmission between new hosts. This transmission potential can be seen as the 'epidemiological' viral fitness and in epidemiological theory is assimilated to the basic reproductive value R_0 of the virus. In simple terms, R_0 represents the average number of secondary infections produced from an infected host in a population of susceptible ones (Fig. 14.4). If $R_0 > 1$, then the virus will become epidemic. By contrast, if $R_0 < 1$,

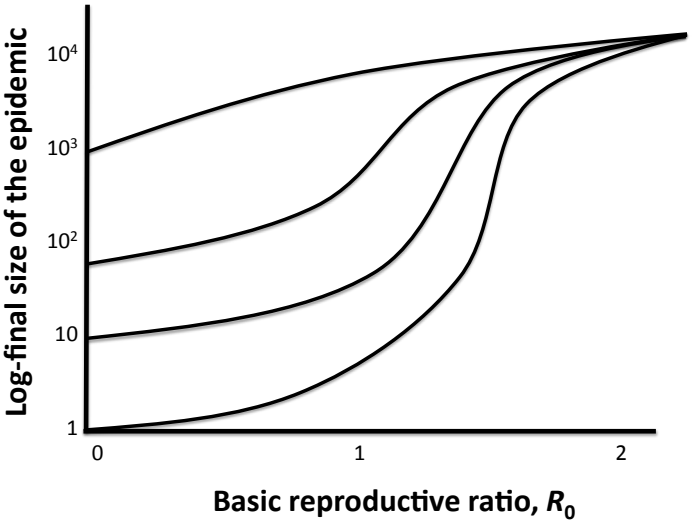

Figure 14.4 Effect of the basic reproductive rate, R_0, of an emerging virus on the size of the epidemic produced. The different curves represent different values for the number of initial infections, I_0. The more initially infected individuals, the less steep the curve. The recursion equation relating these two variables with the final size of the epidemic, I_f, is: $I_f = N - (N - I_0)\exp(-R_0 I_f/N)$, where N represents the size of the susceptible population (Woolhouse *et al.* 2001).

the virus is not transmitted successfully enough to produce a large epidemic, and eventually the virus will disappear from the host population.

Cross-species continuous introduction versus host-to-host transmission

The final number of infected individuals can also be increased in two ways. In the first scenario (Fig. 14.5a), the emerging virus is only accidentally introduced into a local population of the new host for which the contact with the reservoir host is only sporadic. If the infected hosts are moved quickly, or the transmission vector does so, the chances of transmission from new host to new host increase, a positive feedback loop is established, and then the emerging virus will have an $R_0 > 1$ in the general host population. In the second scenario (Fig. 14.5b), the local host population remains in close contact with the reservoir host, making cross-species jumps very likely events. However, despite the virus having $R_0 < 1$ in the global population (represented in Fig. 14.5b by the narrow vertical arrow), the constant reintroduction of the virus creates many opportunities for secondary transmissions. Therefore, although each individual transmission event maybe condemned to extinction, the continuous spillovers from the reservoir are enough to sustain the virus in the new host population.

The role of genetic variability for susceptibility among hosts

A concept that has been recently coined and is gaining interest among epidemiologists is that of *superspreaders*, defined as infected individuals who passed on the infectious agent to many more new hosts than average (Lloyd-Smith *et al.*, 2005; Yates *et al.*, 2006). The phenomenon of superspreading can be viewed as an extreme case of variation or heterogeneity in epidemiological parameters. Superspreaders have been considered as important for the spread of emerging human diseases such as SARS or HIV-1 (Yates *et al.*, 2006), although their importance in emerging plant viral diseases has not been explored yet. Yates *et al.* (2006) have developed mathematical models to account for host heterogeneity in transmission capacity, ranging from non-transmissors to superspreaders. In short, they found that host heterogeneity in susceptibility does not increase the probability of disease emergence, but to the contrary it should even decrease the rate at which

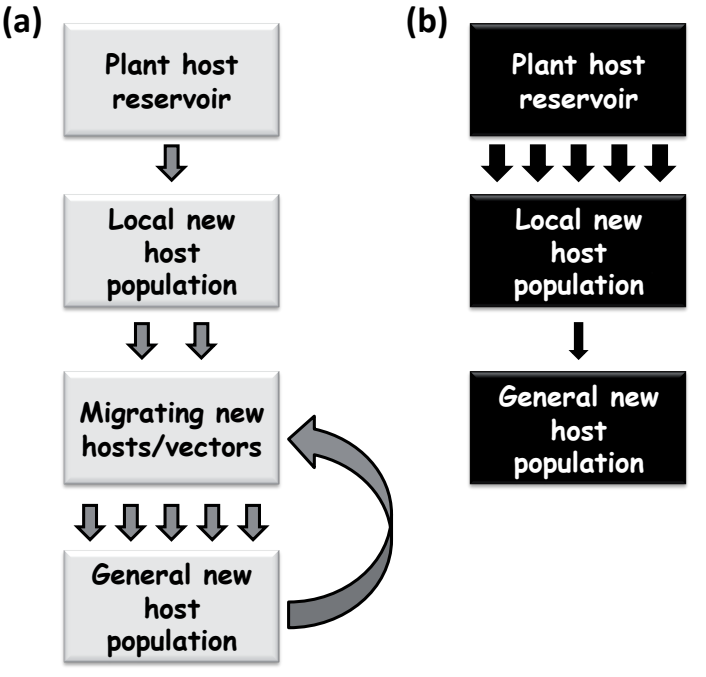

Figure 14.5 Steps in the emergence of an epidemic pathogen. The size of the vertical arrows indicates the likelihood of the event. (a) A situation in which the emerging virus is rarely introduced into the new host population but has the ability to be transmitted among new hosts either by itself or by a vector. (b) The opposite situation, in which the virus recursively spills over from the reservoir into a local population of new hosts, but has little or no capacity for among-hosts transmission. Despite this, the continuous re-introduction facilitates the persistence of the virus in the new host population.

the new virus spreads in the host population, a result that is in good agreement with those produced by other modelling approaches (Lloyd-Smith *et al.*, 2005; Day *et al.*, 2006). Variability in infectivity reduces the risk of emergence. However, simultaneous variability in both traits generates complex results; for example, while variation in susceptibility alone gives the same effect as a homogeneous population with the same R_0, when combined with heterogeneity in mixing, it reduces the risk of emergence compared to the homogeneous case (Yates *et al.*, 2006). Furthermore, Regoes *et al.* (2000) predicted that the host's genetic variability for susceptibility prevents virulence increasing without bounds, which would lead to the evolution of generalist viral strains.

Some of these theoretical predictions have been experimentally validated. For instance, using bacteriophage SBW25Φ2 and mixtures of susceptible and non-susceptible strains of *Pseudomonas fluorescens*, Benmayor *et al.* (2009) have recently shown that an increase in the frequency of the susceptible hosts in the population has two opposing effects: on the one hand, an excess of susceptible hosts allows for mutant viruses with improved performance in the non-susceptible host to appear and rise in frequency. On the other hand, an excess of susceptible hosts reduces the intensity of selection for infecting non-susceptible host genotypes. Therefore, experimental results suggest that the probability of disease emergence is maximal at intermediate frequencies of the susceptible host genotypes.

The role of vector transmission: more bottlenecks

Transmission events, especially when mediated by vectors, such as insects, add a layer of complexity to the emergence process. It is not my aim to provide an exhaustive review on how virus and vectors interact (see Chapter 5). Obviously, for viruses that are transmitted in a persistent and replicative manner, the vector itself represents a host for the virus and the fitness trade-offs described above will apply and contribute to restrict the capacity of the virus to adapt to its plant host. For those viruses that are transmitted in a non-persistent non-replicative manner, the vector is equivalent to a syringe and the constraints that it may impose may be minor, although obviously not null, since the right interaction between viral and insect proteins should be required for successful transmission (Uzest *et al.*, 2007).

Regardless of whether transmission involves replication in the vector or not, a common feature of vector transmission is that it imposes a bottleneck on the virus population, and beneficial variants that appeared in a plant may be lost during the transmission process simply by chance. The question that needs to be answered then is how important is the bottleneck during horizontal transmission? Several studies have tackled this problem experimentally. Ali *et al.* (2006) determined that the bottleneck imposed on horizontal non-persistent transmission of CMV by two different vector species, *Aphis gossypii* and *Myzus persicae*, was strong. Interestingly, these authors found that most of the genetic variability present in the CMV donor population was not lost during the phase of acquisition by the insect, but during the subsequent inoculation phase. As I did above for computing the N_e associated with systemic movement, I have also now applied the variance components method to calculate the expected bottleneck size from the data reported by Ali *et al.* (2006). The estimate, which was robust across experimental blocks and for both aphid species, ranged between 1 and 14 infectious particles transmitted per aphid. In another experiment also involving CMV and *A. gossypii*, Betancourt *et al.* (2008) estimated that the bottleneck size was between one and two viral particles transmitted per aphid, in good agreement with the previous study. Finally, Moury *et al.* (2007) also estimated N_e for the transmission of PVY by *M. persicae* among tobacco plants. These authors estimated a value in the lower range of those reported for CMV: between 1 and 3 viral particles per insect. Personally, I find a bit intriguing that two viruses as different as CMV and PVY have such a similar values of N_e. Naïvely, I would expect the tripartite genome of CMV to be more difficult to transmit than the monopartite genome of PVY, thus producing a lower N_e. Certainly, this theoretical disadvantage for transmission of CMV could be compensated by the differences in transmission strategy: CMV coat protein (CP) interacts directly with the stylet receptor (Chen and Francki 1990), whereas the interaction between PVY CP and the stylet receptor is mediated by HC-Pro (Blanc *et al.*, 1997). Not to mention that the receptors used by each virus may be different or that differences may rise from the fact that in Moury *et al.* (2007) the aphids acquired the virus from an artificial feeding solution, whereas in both CMV studies the aphids feed on infected leaves.

Obviously, the strong bottlenecks associated with transmission by a single insect discussed in the previous paragraph may have no relevance at all in an ecological context because the drift effect may be overcompensated by the population size of the vector aphid and its mobility.

Other relevant host demographic parameters: population size, metapopulation structure

I do not want to close the discussion on Phase III without mentioning, even briefly, two more factors that may contribute to the epidemic spread of an emerging plant virus: the population size of the new host and its spatial distribution. Clearly, the larger the host population size and the more connected, the easier for the virus to spread. By contrast, small and isolated populations would not allow for epidemic spread.

Conclusions

Most of the material I brought together for this chapter explores the role played by viral evolution in the process of emergence. I would like to argue here that the viral genetic variability contained in

the reservoir population is the most important genetic determinant of viral emergence. Natural selection will operate upon this genetic variability to optimize viral fitness during phase II. After reading the discussion presented regarding phases I and II, one may consider that successful emergence, characterized by sustained host-to-host transmission, may be a far more difficult process than might be expected given the remarkable evolutionary plasticity of RNA viruses. Fitness trade-offs, pleiotropic fitness effects, strong bottlenecks at different levels, an excess of deleterious mutations, spatial constraints, fragmented host populations… all together will limit the chances of new viruses to emerge. Therefore, the emergent viruses that we are witnessing nowadays may just represent the few lucky cases that have been able to surmount all these limitations. Our farm animals, crops and we are fortunate that all these limitations exist. Otherwise the number of emerging viruses would be far greater.

The unpredictability of virus emergence means that the only defence we may have for now is to identify and monitor crops at high-risk locations such as tropical deforested regions, places with intense trading activity (specially those suspected of trafficking with illegal plant materials), locations for which changes in the vector fauna are occurring, and places with extensive monocultures in which the spread of a putative emerging virus may be fast. And of course we must keep investing in research that seeks a better understanding of virus evolution. It will be particularly important to pursue ecogenomic projects aimed to catalogue all the many asymptomatic virus infections in wild plants (see Chapter 16) surrounding cultivated areas and that may be important for future cases of emergence.

References

Agudelo-Romero, P., Carbonell, P., de la Iglesia, F., Carrera, J., Rodrigo, G., Jaramillo, A., Pérez-Amador, M.A., and Elena, S.F. (2008a). Changes in gene expression profile of *Arabidopsis thaliana* after infection with tobacco etch virus. Virol. J. *5*, 92.

Agudelo-Romero, P., Carbonell, P., Pérez-Amador, M.A., and Elena, S.F. (2008b). Virus adaptation by manipulation of host's gene expression. PLoS ONE *3*, e2397.

Ali, A., Li, H., Schneider, W.L., Sherman, D.J., Gray, S., Smith, D., and Roossinck, M.J. (2006). Analysis of genetic bottlenecks during horizontal transmission of cucumber mosaic virus. J. Virol. *17*, 8345–8350.

Anderson, P.K., Cunningham, A.A., Patel, N.G., Morales, F.J., Epstein, P.R., and Daszak, P. (2004). Emerg. Infect. Dis. of plants: pathogen pollution, climate change and agrotechnology drivers. Trends Ecol. Evol. *19*, 535–544.

Benmayor, R., Hodgson, D.J., Perron, G.G., and Buckling, A. (2009). Host mixing and disease emergence. Curr. Biol. *19*, 764–767.

Betancourt, M., Fereres, A., Fraile, A., and García-Arenal, F. (2008). Estimation of the effective number of founders that initiate an infection after aphid transmission of a multipartite plant virus. J. Virol. *82*, 12416–12421.

Blanc, S., López-Moya, J.J., Wang, R.Y., García-Lampasona, S., Thornbury, D.W., and Pirone, T.P. (1997). A specific interaction between coat protein and helper component correlates with aphid transmission of a potyvirus. Virology *231*, 141–147.

Bonhoeffer, S., Chappey, C., Parkin, N.T., Whitcomb, J.M., and Petropoulos, C.J. (2004). Evidence for positive epistasis in HIV-1. Science *306*, 1547–1550.

Burch, C.L., and Chao, L. (2004). Epistasis and its relationship to canalization in the RNA virus φ6. Genetics *167*, 559–567.

Carrasco, P., de la Iglesia, F., and Elena, S.F. (2007). Distribution of fitness and virulence effects caused by single-nucleotide substitutions in tobacco etch virus. J. Virol. *81*, 12979–12984.

Chare, E.R., Gould, E.A., and Holmes, E.C. (2003). Phylogenetic analysis reveals a low rate of homologous recombination in negative-sense RNA viruses. J. Gen. Virol. *85*, 3149–3157.

Chare, E.R., and Holmes, E.C. (2006). A phylogenetic survey of recombination frequency in plant RNA viruses. Arch. Virol. *151*, 933–946.

Chen, B., and Francki, R.I.B. (1990). Cucumovirus transmission by the aphids *Myzus persicae* is determined solely by the coat protein. J. Gen. Virol. *71*, 939–944.

Cleaveland, S., Haydon, D.T., and Taylor, L. (2007). Overviews of pathogen emergence: Which pathogens emerge, when and why? Curr. Top. Microbiol. Immunol. *315*, 85–111.

Codoñer, F.M., and Elena, S.F. (2008). The promiscuous evolutionary history of the family *Bromoviridae*. J. Gen. Virol. *89*, 1739–1747.

Day, T., André, J.B., and Park, A. (2006). The evolutionary emergence of pandemic influenza. Proc. R. Soc. Lond. B 273, 2945–2953.

DeFilippis, V.R., and Villareal, L.P. (2000). An introduction to the evolutionary ecology of viruses. In Viral Ecology, Hurst, C.J., ed. (New York: Academic Press), pp. 126–208.

Díaz-Pendón, J.A., and Ding, S.W. (2008). Direct and indirect roles of viral suppressors of RNA silencing in pathogenesis. Annu. Rev. Phytopathol. *46*, 303–326.

Dietrich, C., and Maiss, E. (2003). Fluorescent labeling reveals spatial separation of potyvirus populations in mixed infected *Nicotiana bethamiana* plants. J. Gen. Virol. *84*, 2871–2876.

Domingo-Calap, P., Cuevas, J.M., and Sanjuán, R. (2009). The fitness effects of random mutations in single-stranded DNA and RNA bacteriophages. PLoS Genet. *5*, e1000742.

Dunoyer, P., and Voinnet, O. (2005). The complex interplay between plant viruses and host RNA silencing pathways. Curr. Op. Plant Biol. *8*, 415–423.

Elena, S.F., Agudelo-Romero, P., and Lalic, J. (2009). The evolution of viruses in multi-host fitness landscapes. Open Virol. J. *3*, 1–6.

Elena, S.F., Carrasco, P., Daròs, J.A., and Sanjuán, R. (2006). Mechanisms of genetic robustness in RNA viruses. EMBO Rep. *7*, 168–173.

Elena, S.F., and Sanjuán, R. (2008). Virus evolution: insights from an experimental approach. Annu. Rev. Ecol. Evol. Syst. *38*, 27–52.

Escriu, F., Fraile, A., and García-Arenal, F. (2000). Evolution of virulence in natural populations of the satellite RNA of cucumber mosaic virus. Phytophatology *90*, 480–495.

Frank, S.A. (2001). Multiplicity of infection and the evolution of hybrid imcompatibilities in segmented viruses. Heredity *87*, 522–529.

Froissart, R., Roze, D., Uzest, M., Galibiert, L., Blanc, S., and Michalakis, Y. (2005). Recombination every day: abundant recombination in a virus during a single multi-cellular host infection. PLoS Biol. *3*, e89.

Gago, S., Elena, S.F., Flores, R., and Sanjuán, R. (2009). Extremely high mutation rate of a hammerhead viroid. Science *323*, 1308.

García-Cano, E., Resende, R.O., Fernández-Muñoz, R., and Moriones, E. (2006). Synergistic interaction between tomato chlorosis virus and tomato spotted wilt virus results in breakdown of resistance in tomato. Phytopathology *96*, 1263–1269.

Garrett, K.A., Dendy, S.P., Frank, E.E., Rouse, M.N., and Travers, S.E. (2006). Climate change effects on plant disease: genomes to ecosystems. Annu. Rev. Phytopathol. *44*, 489–509.

González-Jara, P., Fraile, A., Cantó, T., and García-Arenal, F. (2009). The multiplicity of infection of a plant virus varies during colonization of its eukaryotic host. J. Virol. *83*, 7487–7494.

Guerini, M.N., and Murphy, J.F. (1999). Resistance of *Capsicum annuum* 'Avelar' to pepper mottle potyvirus and alleviation of this resistance by co-infection with cucumber mosaic cucumovirus are associated with virus movement. J. Gen. Virol. *80*, 2785–2792.

Hacker, D.L., and Fowler, B.C. (2000). Complementation of the host range restriction of southern cowpea mosaic virus in bean by southern bean mosaic virus. Virology *266*, 140–149.

Hall, J.S., French, R., Hein, G.L., Morris, T.J., and Stenger, D.C. (2001a). Three distinct mechanisms facilitate genetic isolation of sympatric wheat streak mosaic virus lineages. Virology *282*, 230–236.

Hall, J.S., French, R., Morris, T.J., and Stenger, D.C. (2001b). Structure and temporal dynamics of populations within wheat streak mosaic virus isolates. J. Virol. *75*, 10231–10243.

Hammond, J., Lecoq, H., and Raccah, B. (1999). Epidemiological risks from mixed virus infections and transgenic plants expressing viral genes. Adv. Virus Res. *54*, 189–314.

Holmes, E.C. (2008). The evolutionary history and phylogeography of human viruses. Annu. Rev. Microbiol. *62*, 307–328.

Holmes, E.C., and Drummond, A.J. (2007). The evolutionary genetics of viral emergence. Curr. Top. Microbiol. Immunol. *315*, 51–66.

Jones, R.A.C. (2009). Plant virus emergence and evolution: origins, new encounter scenarios, factors driving emergence, effects of changing world conditions, and prospects for control. Virus Res. *141*, 113–130.

Jones, J.D.G., and Dangl, J.L. (2006). The plant immune system. Nature *444*, 323–329.

Jridi, C., Martin, J.F., Mareie-Jeanne, V., Labonne, G., and Blanc, S. (2006). Distinct viral populations differentiate and evolve independently in a single perennial host plant. J. Virol. *80*, 2349–2357.

Kearney, C.M., Thomson, M.J., and Roland, K.E. (1999). Genome evolution of tobacco mosaic virus populations during long-term passaging in a diverse range of hosts. Arch. Virol. *144*, 1513–1526.

Király, L., Barna, B., and Király, Z. (2007). Plant resistance to pathogen infection: forms and mechanisms of innate and acquired resistance. J. Phytopathol. *155*, 385–396.

Kokkinos, C.D., and Clark, C.A. (2006). Interactions among sweet potato chlorotic stunt virus and different potyviruses and potyvirus strains infecting sweetpotato in the United States. Plant Dis. *90*, 1347–1352.

Li, F., and Ding, S.W. (2006). Virus counterdefense: diverse strategies for evading the RNA-silencing immunity. Annu. Rev. Microbiol. *60*, 503–531.

Li, H., and Roossinck, M.J. (2004). Genetic bottlenecks reduce population variation in an experimental RNA virus population. J. Virol. *78*, 10582–10587.

Lin, S.S., Wu, H.W., Elena, S.F., Chen, K.C., Niu, Q.W., Yeh, S.D., Chen, C.C., and Chua, N.H. (2009). Molecular evolution of a viral non-coding sequence under the selective pressure of amiRNA-mediated silencing. PLoS Pathog. *5*, e1000312.

Lloyd-Smith, J.O., Schreiber, S.J., Kopp, P.E., and Getz, W.M. (2005). Superspreading and the effect of individual variation on disease emergence. Nature *438*, 355–359.

Malpica, J.M., Fraile, A., Moreno, I., Obies, C.I., Drake, J.W., and García-Arenal, F. (2002). The rate and character of spontaneous mutation in an RNA virus. Genetics *162*, 1505–1511.

Malpica, J.M., Sacristán, S., Fraile, A., and García-Arenal, F. (2006). Association and host selectivity in multihost pathogens. PLoS ONE *1*, e41.

Martin, G., and Lenormand, T. (2006). The fitness effect of mutations across environments: a survey in light of fitness landscapes models. Evolution *60*, 2413–2427.

Martín, S., and Elena, S.F. (2009). Application of game theory to the interaction between plant viruses during mixed infections. J. Gen. Virol. *90*, 2815–2820.

Monsion, B., Froissart, R., Michalakis, Y., and Blanc, S. (2008). Large bottleneck size in cauliflower mosaic virus populations during host plant colonization. PLoS Pathog. *4*, e1000174.

Moreno, I., Malpica, J.M., Rodríguez-Cerezo, E., and García-Arenal, F. (1997). A mutation in tomato aspermy cucumovirus that abolishes cell-to-cell movement is maintained to high levels in the viral RNA population by complementation. J. Virol. *71*, 9157–9162.

Moury, B., Fabre, F., and Senoussi, R. (2007). Estimation of the number of virus particles transmitted by an insect vector. Proc. Natl. Acad. Sci. U.S.A. *45*, 17891–17896.

Parrish, C.R., Holmes, E.C., Morens, D.M., Park, E.C., Burke, D.S., Calisher, C.H., Laughlin, C.A., Saif, L.J., and Daszak, P. (2008). Cross-species virus transmission and the emergence of new epidemic diseases. Microbiol. Mol. Biol. Rev. *72*, 457–470.

Pita, J.S., De Miranda, J.R., Schneider, W.L., and Roossinck, M.J. (2007). Environment determines fidelity for an RNA virus replicase. J. Virol. *81*, 9072–9077.

Regoes, R.R., Nowak, M.A., and Bonhoeffer, S. (2000). Evolution of virulence in a heterogeneous host population. Evolution *54*, 64–71.

Rokyta, D.R., Joyce, P., Caudle, S.B., and Wichman, H.A. (2005). An empirical test of the mutational landscape model of adaptation using a single-stranded DNA virus. Nat. Genet. *37*, 441–444.

Sacristán, S., Malpica, J.M., Fraile, A., and García-Arenal, F. (2003). Estimation of population bottlenecks during systemic movement of *tobacco mosaic virus* in tobacco plants. J. Virol. *77*, 9906–9911.

Sánchez-Navarro, J.A., Herranz, M.C., and Pallás, V. (2006). Cell-to-cell movement of alfalfa mosaic virus can be mediated by the movement proteins of ilar-, bromo-, cucumo-, tobamo- and comoviruses *and* does not require virion formation. Virology *346*, 66–73.

Sanjuán, R., Agudelo-Romero, P., and Elena, S.F. (2009). Upper-limit mutation rate estimation for a plant RNA virus. Biol. Lett. *5*, 394–396.

Sanjuán, R., Cuevas, J.M., Moya, A., and Elena, S.F. (2005). Epistasis and the adaptability of an RNA virus. Genetics *170*, 1001–1008.

Sanjuán, R., Moya, A., and Elena, S.F. (2004a). The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus. Proc. Natl. Acad. Sci. U.S.A. *101*, 8396–8401.

Sanjuán, R., Moya, A., and Elena, S.F. (2004b). The contribution of epistasis to the architecture of fitness in an RNA virus. Proc. Natl. Acad. Sci. U.S.A. *101*, 15376–15379.

Sardanyés, J., Solé, R.V., and Elena, S.F. (2009). Replication mode and landscape topology differentially affect RNA virus mutational load and robustness. J. Virol. *83*, 12579–12589.

Schneider, W.L., and Roossinck, M.J. (2000). Evolutionarily related Sindbis-like plant viruses maintain different levels of population diversity in a common host. J. Virol. *74*, 3130–3134.

Schneider, W.L., and Roossinck, M.J. (2001). Genetic diversity in RNA virus quasispecies is controlled by host–virus interactions. J. Virol. *75*, 6566–6571.

Thresh, J.M. (2006). Plant virus epidemiology: the concept of host genetic vulnerability. Adv. Virus Res. *67*, 89–125.

Torres-Barceló, C., Daròs, J.A., and Elena, S.F. (2009). Compensatory molecular evolution of HC-Pro, an RNA-silencing suppressor from a plant RNA virus. Mol. Biol. Evol. DOI: 10.1093/molbev/msp272.

Uzest, M., Gargani, D., Drucker, M., Hébrard, E., Garzo, E., Candresse, T., Fereres, A., and Blanc, S. (2007). A protein key to plant virus transmission at the tip of the insect vector stylet. Proc. Natl. Acad. Sci. U.S.A. *104*, 17959–17964.

Van der Vlugt, R.A.C., Stijger, C.M., Verhoeven, J.T.J., and Lesemann, D.E. (2000). First report of pepino mosaic virus in tomato. Plant Dis. *84*, 103–108.

Voinnet, O. (2001). RNA silencing as a plant immune system against viruses. Trends Genet. *17*, 449–459.

Waterhouse, P.M., Wang, M.B., and Lough, T. (2001). Gene silencing as an adaptive defence against viruses. Nature *411*, 834–842.

Weinreich, D.M., Watson, R.A., and Chao, L. (2005). Sign epistasis and genetic constraints on evolutionary trajectories. Evolution *59*, 1165–1174.

Wilke, C.O., Forster, R., and Novella, I.S. (2006). Quasispecies in time-dependent environments. Curr. Top. Microbiol. Immunol *299*, 33–50.

Wintermantel, W.M., Cortez, A.A., Anchieta, A.G., Gulati-Sakhuja, A., and Hladky, L.L. (2008). Co-infection by two criniviruses alters accumulation of each virus in a host-specific manner and influences efficiency of virus transmission. Phytopathology *98*, 1340–1345.

Woolhouse, M.E.J., and Dye, C. (2001). Population biology of emerging and reemerging pathogens – Preface. Phil. Trans. R. Soc. Lond. B 356, 981–982.

Woolhouse, M.E.J., Taylor, L.H., and Haydon, D.T. (2001). Population biology of multihost pathogens. Science *292*, 1109–1112.

Woolhouse, M.E.J., Haydon, D.T., and Antia, R. (2005). Emerging pathogens: the epidemiology and evolution of species jumps. Trends Ecol. Evol. *20*, 238–244.

Yates, A., Antia, R., and Regoes, R.R. (2006). How do pathogen evolution and host heterogeneity interact in disease emergence? Proc. R. Soc. B. *273*, 3075–3083.